09/914286
518 Re PCT/PTO 2 4 AUG 2001

SPECIFICATION

2/prts

insert a1

Avermectin Aglycon Synthase Genes

Technical Field

The present invention relates to DNAs encoding multifunctional enzyme involved in the biosynthesis of an avermectin compound which is a polyketide; polypeptides encoded by the DNAs; vectors containing the DNAs; host cells transformed with the DNAs or the vectors; and a process for producing avermectin.

Background Art

A polyketide is a group of compounds containing a number of natural substances which vary in their structures and functions. Polyketides are known to include compounds having a variety of bioactivities such as antibacterial agents, antimyotic agents, antiparasitic agents, anti-insect agents, antitumor agents, and immunosuppressant agents, and aromatic compounds which are produced by bacteria, fungi and plants.

The above-mentioned various polyketide compounds are synthesized by the same biosynthetic mechanism which is very similar to the biosynthesis of fatty acids. That is, a polyketide compound is biosynthesized by the steps of continuous condensation of lower fatty acids including acetic acids and propionic acids, and subsequent reactions such as reduction of ketone, dehydration and enoyl reduction of each carbonyl group at β position of the extended acyl group which is similar to fatty acid synthesis. These various repetitive synthetic processes of many polyketide compounds are carried out a macromolecule, multifunctional enzyme complex, which has specific active sites (domains) required for each catalytic activity. A general reaction manner of polyketide biosynthesis is outlined, for example in Ann. Rev. Gen., 24, 37 (1990), and Ann. Rev. Microbiol., 47, 875 (1993).

It has been shown that a DNA sequence encoding polyketide synthase usually encodes all the required activities for the synthesis of a polyketide backbone. The DNA sequence encoding polyketide synthase is composed of modules, that is, repeating units involving condensation steps and modification steps following condensation. Each catalytic activity is involved in specificity to a specific carboxylic acid component of each condensation step, or in a different site which specifies a modification function following a specific condensation step to be achieved. For example, International Publication WO93/13663 describes the constitution of a gene encoding polyketide synthase of *Saccharopolyspora erythaea*. This gene consists of 6 modules, each of which is responsible for one condensation step. That is, a correct sequence of acyl side chain elongation and modification of an elongating chain are determined by genetic information present in each module.

Regarding the biosynthetic mechanism of avermectin aglycon, it has been reported that like other polyketide compounds, synthesis units of avermectin aglycon are lower fatty acids, such as acetic acid and propionic acid as its components [J. Antibiot., 39, 541-549 (1986)], and as in *Saccharopolyspora erythaea*, polyketide synthase consisting of modules is present in avermectin-producing bacteria [Gene, 115, 119-125 (1992), Ann. New York Acad. of Sci., 721, 123-132 (1994)].

Japanese Published Unexamined Patent Application No. 15391/91 describes a DNA fragment involved in avermectin biosynthesis, but shows no nucleotide sequence of the DNA fragment. This publication merely assumes the presence of polyketide synthase, which is involved in the synthesis of avermectin aglycon and the presence of partial modules. Therefore, the entire structure of polyketide synthase of avermectin cannot be predicted.

Similarly, MacNeil et al have reported a domain structure of partial modules

[Ann. New York Acad. of Sci., 721, 123-132 (1994)]. However, they have not revealed the nucleotide sequence that should be evidence for polyketide synthase of avermectin.

Alteration of polyketide synthase would be a very useful breeding technique upon breeding of bacterial strains which can be used for a novel process for producing a novel avermectin useful as veterinary drugs and agricultural chemicals, and can produce a more effective avermectin derivative. Steps required to carry out such alteration include determination of the entire nucleotide sequence of a gene encoding polyketide synthase, accurate determination of a domain structure of each module based on the sequence, and introduction of a desired mutation. However, as described above, it was very difficult to carry out such improved-breeding, since the polyketide synthase gene of avermectin aglycon had not been specified and the nucleotide sequence of the gene was unknown.

The present inventors have studied approaches for producing a component different from that produced by the wild type strain by engineering DNA involved in polyketide synthesis with various methods. To apply this methodology, first we had to isolate a DNA molecule involved in the biosynthesis of a polyketide compound.

Hence, an object of the present invention is to provide a DNA encoding a multifunctional enzyme involved in biosynthesis of avermectin aglycon, and a process for producing avermectin aglycon, altered avermectin aglycon, avermectin, and altered avermectin using the DNA.

Disclosure of the Invention

The present inventors made an intensive investigation to attain the object. As a result, the inventors have succeeded in isolating DNAs encoding a multifunctional enzyme involved in biosynthesis of avermectin aglycon. The present invention has

been completed on the basis of this result.

The present invention relates to the following (1) to (43).

(1) A DNA encoding avermectin aglycon synthase (hereinafter, also referred to as an avermectin aglycon synthase gene).

In an embodiment of the present invention, the DNA is derived from a wild-type avermectin-producing strain or a mutant strain thereof, such as one belonging to the genus *Streptomyces*, specifically *Streptomyces avermitilis*.

(2) A DNA comprising a nucleotide sequence selected from the group consisting of nucleotide Nos. 1-11916 and 11971-30688 of SEQ ID NO: 1, and nucleotide Nos. 1-14643 and 14824-31419 of SEQ ID NO: 2; or

a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having avermectin aglycon synthase activity.

The above term "a DNA which hybridizes with this DNA under stringent conditions" refers to a DNA which is obtained by colony hybridization, plaque hybridization or Southern hybridization or the like using the DNA having a nucleotide sequence of SEQ ID NO. 1 or 2. For example, such a DNA can be identified by carrying out hybridization at 65°C in the presence of 0.7 to 1.0 mol/l sodium chloride using a filter on which DNAs derived from colonies or plaques have been immobilized, followed by washing the filter at 65°C using 0.1 to 2-fold concentrated SSC solution (1-fold concentrated SSC solution consists of 150 mmol/l sodium chloride, 15 mmol/l sodium citrate).

Hybridization can be carried out according to a method described in experimental protocols, such as Molecular Cloning, A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, 1989 (hereinafter abbreviated as

Molecular Cloning 2nd Edition), Current Protocols in Molecular Biology, John Wiley & Sons, 1987-1997 (hereinafter abbreviated as Current Protocols in Molecular Biology), DNA Cloning 1; Core Techniques, A Practical Approach, Second Edition, Oxford University, 1995. Specific examples of the DNA which can be hybridized include a DNA having at least homology of 80% or more, preferably 95% or more with a nucleotide sequence selected from the group consisting of nucleotide Nos. 1-11916 and 11971-30688 of SEQ ID NO: 1, and nucleotide Nos. 1-14643 and 14824-31419 of SEQ ID NO: 2.

The following term "a DNA which hybridizes with this DNA (or said DNA) under stringent conditions" can also be defined in the same manner as described above.

(3) The DNA according to the above (1) or (2) wherein the DNA comprises DNAs encoding avermectin aglycon synthase domains.

(4) The DNA according to the above (3) wherein the DNA encoding avermectin aglycon synthase domains is selected from the group consisting of:

a DNA encoding a polypeptide having acyltransferase activity and acyl carrier protein activity;

a DNA encoding a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, β-ketoacyl-ACP reductase activity and acyl carrier protein activity;

a DNA encoding a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, dehydratase activity, β-ketoacyl-ACP reductase activity and acyl carrier protein activity;

a DNA encoding a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, and acyl carrier protein activity; and

a DNA encoding a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, dehydratase activity, β-ketoacyl-ACP reductase activity, acyl

carrier protein activity, and thioesterase activity.

(5) The DNA according to the above (4) wherein the DNA encoding a polypeptide having acyltransferase activity and acyl carrier protein activity is a DNA comprising the nucleotide sequence of nucleotide Nos. 85-1353 of SEQ ID NO: 1; or
a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having acyltransferase activity and acyl carrier protein activity.

(6) The DNA according to the above (4) wherein the DNA encoding a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, β-ketoacyl-ACP reductase activity, and acyl carrier protein activity is:
a DNA comprising the nucleotide sequence selected from the group consisting of nucleotide Nos. 1441-6180, 15217-19938 and 20008-24690 of SEQ ID NO: 1, and nucleotide Nos. 100-4692 and 14935-20334 of SEQ ID NO: 2; or
a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, β-ketoacyl-ACP reductase activity, and acyl carrier protein activity.

(7) The DNA according to the above (4) wherein the DNA encoding a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, dehydratase activity, β-ketoacyl-ACP reductase activity, and acyl carrier protein activity is:
a DNA comprising the nucleotide sequence selected from the group consisting of nucleotide Nos. 6256-11658 and 24781-30309 of SEQ ID NO: 1, and nucleotide Nos. 20413-25734 and 25810-31125 of SEQ ID NO: 2; or
a DNA which hybridizes with this DNA under stringent conditions and which encodes polypeptides having β-ketoacyl-ACP synthase activity, acyltransferase activity, dehydratase activity, β-ketoacyl-ACP reductase activity, and acyl carrier protein activity.

(8) The DNA according to the above (4) wherein the DNA encoding a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, and acyl carrier protein activity is:

a DNA comprising the nucleotide sequence of nucleotide No. 12076-15147 of SEQ ID NO: 1, or nucleotide No. 4771-7818 of SEQ ID NO: 2;

or a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, and acyl carrier protein activity.

(9) The DNA according to the above (4) wherein the DNA encoding a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, dehydratase activity, β-ketoacyl-ACP reductase activity, acyl carrier protein activity, and thioesterase activity is:

a DNA comprising the nucleotide sequence of nucleotide Nos. 7906-14619 of SEQ ID NO: 2; or

a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, dehydratase activity, β-ketoacyl-ACP reductase activity, acyl carrier protein activity, and thioesterase activity.

(10) The DNA according to the above (4) wherein the DNA encoding a polypeptide having acyltransferase activity is:

a DNA comprising the nucleotide sequence selected from the group consisting of nucleotide Nos. 85-1032, 7906-8829, 13756-14694, 16917-17862, 21658-22584, and 26413-27336 of SEQ ID NO: 1, and nucleotide Nos. 1648-2673, 6322-7344, 9676-10773, 16543-17565, 21991-23019 and 27367-28392 of SEQ ID NO: 2; or

a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having acyltransferase activity.

(11) The DNA according to the above (4) wherein the DNA encoding a polypeptide having acyl carrier protein activity is:

a DNA comprising the nucleotide sequence selected from the group consisting of nucleotide Nos. 1096-1353, 5935-6180, 11413-11658, 14902-15147, 19693-19938, 24445-24690 and 30064-30309 of SEQ ID NO: 1, and nucleotide Nos. 4447-4692, 7573-7818, 13378-13659, 20089-20334, 25489-25734 and 30880-31125 of SEQ ID NO: 2; or

a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having acyl carrier protein activity.

Sub a2 (12) The DNA according to the above (4) wherein the DNA encoding a polypeptide having β-ketoacyl-ACP synthase activity is:

a DNA comprising the nucleotide sequence selected from the group consisting of nucleotide Nos. 1441-2742, 6256-7545, 12076-13368, 15217-16506, 20008-21297 and 24781-26079 of SEQ ID NO: 1, and nucleotide Nos. 100-1383, 4771-6060, 7906-9258, 14935-16224, 20413-21705 and 25810-27102 of SEQ ID NO: 1; or

a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having β-ketoacyl-ACP synthase activity.

(13) The DNA according to the above (4) wherein the DNA encoding a polypeptide having β-ketoacyl-ACP reductase activity is:

a DNA comprising the nucleotide sequence selected from the group consisting of nucleotide Nos. 5143-5676, 10609-11142, 18886-19419, 23602-24138 and 29227-29760 of SEQ ID NO: 1, and nucleotide Nos. 3634-4188, 12547-13104, 19285-19842, 24685-25242 and 30076-30633 of SEQ ID NO: 2; or

a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having β-ketoacyl-ACP reductase activity.

(14) The DNA according to the above (4) wherein the DNA encoding a

polypeptide having dehydratase activity is:

a DNA comprising the nucleotide sequence selected from the group consisting of nucleotide Nos. 8947-9384 and 27475-27894 of SEQ ID NO: 1, and nucleotide Nos. 10885-11289, 23149-23529 and 28516-28878 of SEQ ID NO: 2; or

a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having dehydratase activity.

(15) The DNA according to the above (4) wherein the DNA encoding a polypeptide having thioesterase activity is:

a DNA having the nucleotide sequence of nucleotide No. 13879-14619 of SEQ ID NO: 2; or

a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having thioesterase activity.

(16) The DNA according to the above (3) or (4) wherein the DNA encoding an avermectin aglycon synthase domain is a mutated DNA encoding a polypeptide having enhanced or diminished activity of the domain.

(17) The DNA according to the above (16) wherein the DNA encoding a polypeptide having diminished activity of avermectin aglycon synthase domain is the DNA comprising a nucleotide sequence of SEQ ID NO: 7.

(18) A DNA encoding an avermectin aglycon synthase domain which comprises a nucleotide sequence selected from the group consisting of nucleotide Nos. 85-1032, 1096-1353, 1441-2742, 3148-4068, 5143-5676, 5935-6180, 6256-7545, 7906-8829, 8947-9384, 10609-11142, 11413-11658, 12076-13368, 13756-14694, 14902-15147, 15217-16506, 16917-17862, 18886-19419, 19693-19938, 20008-21297, 21658-22584, 23602-24138, 24445-24690, 24781-26079, 26413-27336, 27475-27894, 29227-29760 and 30064-30309 of SEQ ID NO: 1, and nucleotide Nos. 100-1383, 1648-2673,

3634-4188, 4447-4692, 4771-6060, 6322-7344, 7573-7818, 7906-9258, 9676-10773, 10885-11289, 12547-13104, 13378-13659, 13879-14619, 14935-16224, 16543-17565, 17689-18066, 19285-19842, 20089-20334, 20413-21705, 21991-23019, 23149-23529, 24685-25242, 25489-25734, 25810-27102, 27367-28392, 28516-28878, 30076-30633, and 30880-31125 of SEQ ID NO: 2; or
a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having avermectin aglycon synthase domain activity.

(19) A DNA comprising the nucleotide sequence of nucleotide No. 85-1353 of SEQ ID NO: 1 ;or
a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having acyltransferase activity and acyl carrier protein activity.

(20) A DNA comprising the nucleotide sequence selected from the group consisting of nucleotide Nos. 1441-6180, 15217-19938 and 20008-24690 of SEQ ID NO: 1, and nucleotide Nos. 100-4692 of SEQ ID NO: 2; or
a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, β-ketoacyl-ACP reductase activity and acyl carrier protein activity.

(21) A DNA comprising the nucleotide sequence selected from the group consisting of nucleotide Nos. 6256-11658 and 24781-30309 of SEQ ID NO: 1, and nucleotide Nos. 14935-20334, 20413-25734 and 25810-31125 of SEQ ID NO: 2; or
a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, dehydratase activity, β-ketoacyl-ACP reductase activity and acyl carrier protein activity.

(22) A DNA comprising the nucleotide sequence of nucleotide No. 12076-15147 of SEQ ID NO: 1, or the nucleotide sequence of nucleotide No. 4771-7818 of SEQ ID

NO: 2; or

a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, and acyl carrier protein activity.

(23) A DNA comprising the nucleotide sequence of nucleotide No. 7906-14619 of SEQ ID NO: 2; or

a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having β-ketoacyl-ACP synthase activity, acyltransferase activity, dehydratase activity, β-ketoacyl-ACP reductase activity, acyl carrier protein activity, and thioesterase activity.

(24) A DNA comprising the nucleotide sequence selected from the group consisting of nucleotide Nos. 85-1032, 7906-8829, 13756-14694, 16917-17862, 21658-22584 and 26413-27336 of SEQ ID NO: 1, and nucleotide Nos. 1648-2673, 6322-7344, 9676-10773, 16543-17565, 21991-23019 and 27367-28392 of SEQ ID NO: 2; or

a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having acyltransferase activity.

(25) A DNA comprising the nucleotide sequence selected from the group consisting of nucleotide Nos. 1096-1353, 5935-6180, 11413-11658, 14902-15147, 19693-19938, 24445-24690, and 30064-30309 of SEQ ID NO: 1, and nucleotide Nos. 4447-4692, 7573-7818, 13378-13659, 20089-20334, 25489-25734 and 30880-31125 of SEQ ID NO: 2; or

a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having acyl carrier protein activity.

Sub a3 (26) A DNA comprising the nucleotide sequence selected from the group

consisting of nucleotide Nos. 1441-2742, 6256-7545, 12076-13368, 15217-16506, 20008-21297 and 24781-26079 of SEQ ID NO: 1, and nucleotide Nos. 100-1383, 4771-6060, 7906-9258, 14935-16224, 20413-21705, and 25810-27102 of SEQ ID NO: 1; or
a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having β-ketoacyl-ACP synthase activity.

(27) A DNA comprising the nucleotide sequence selected from the group consisting of nucleotide Nos. 5143-5676, 10609-11142, 18886-19419, 23602-24138, and 29227-29760 of SEQ ID NO: 1, and nucleotide Nos. 3634-4188, 12547-13104, 19285-19842, 24685-25242 and 30076-30633 of SEQ ID NO: 2; or
a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having β ketoacyl ACP reductase activity.

(28) A DNA comprising the nucleotide sequence selected from the group consisting of nucleotide Nos. 8947-9384 and 27475-27894 of SEQ ID NO: 1, and nucleotide Nos. 10885-11289, 17689-18066, 23149-23529 and 28516-28878 of SEQ ID NO: 2; or
a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having dehydratase activity.

(29) A DNA comprising the nucleotide sequence of nucleotide Nos. 13879-14619 of SEQ ID NO: 2; or
a DNA which hybridizes with this DNA under stringent conditions and which encodes a polypeptide having thioesterase activity.

(30) A DNA comprising the nucleotide sequence shown in SEQ ID NO: 7.

(31) A polypeptide encoded by the DNA according to any one of the above (1) to

(29) .

(32) A polypeptide comprising the amino acid sequence according to any one of SEQ ID NOS: 3 to 6; or

a polypeptide comprising an amino acid sequence wherein one or more amino acids are deleted, replaced or added in the amino acid sequence according to any one of SEQ ID NOS: 3 to 6, and having avermectin aglycon synthase activity.

The above "polypeptide comprising an amino acid sequence wherein one or more amino acids are deleted, replaced or added, and having avermectin aglycon synthase activity" can be prepared according to site-directed mutagenesis as described in Molecular Cloning, 2nd Edition, Current Protocols in Molecular Biology, Nucleic Acids Research, 10, 6487 (1982), Proc. Natl. Acad. Sci., USA, 79, 6409 (1982), Gene, 34, 315 (1985), Nucleic Acids Research, 13, 4431 (1985), Proc. Natl. Acad. Sci USA, 82, 488 (1985) and the like. The number of amino acids which are deleted, replaced or added is not specifically limited, but is a number of amino acids which can be deleted, replaced, or added by known methods, such as the above site-directed mutagenesis, and within the range from 1 to several tens of amino acids, preferably 1 to 20, more preferably 1 to 10, and even more preferably 1 to 5 amino acids.

Sub a4

(33) A polypeptide comprising the amino acid sequence selected from the group consisting of amino acid Nos. 29-344, 366-451, 481-914, 1050-1356, 1715-1892, 1979-2060, 2086-2515, 2983-3128, 3537-3714 and 3805-3886 of SEQ ID NO: 3, amino acid Nos. 36-466, 596-908, 978-1059, 1083-1512, 1653-1964, 2306-2483, 2575-2656, 2680-3109, 32030-3538, 3878-4056, 4159-4240, 4271-4703, 4815-5122, 5753-5930 and 6032-6113 of SEQ ID NO: 4, amino acid Nos. 34-461, 550-891, 1212-1396, 1483-1564, 1591-2020, 2108-2448, 2525-2606, 2636-3086, 3226-3591, 3629-3763, 4183-4363, 4460-4553 and 4627-4873 of SEQ ID NO: 5, amino acid Nos. 38-467, 574-914, 956-1081, 1488-1673, 1756-1837, 1864-2294, 2390-2732, 2776-2902,

3288-3473, 3556-3637, 3663-4093, 4182-4523, 4565-4685, 5085-5270 and 5353-5434 of SEQ ID NO: 6; or

a polypeptide comprising an amino acid sequence wherein one or more amino acids are deleted, replaced or added in the amino acid sequence selected above, and having avermectin aglycon synthase domain activity.

The above "polypeptide comprising an amino acid sequence wherein one or more amino acids are deleted, replaced or added in the amino acid sequence selected above, and having avermectin aglycon synthase domain activity" can be obtained according to the method described in the above (32) .

(34) A recombinant vector comprising the DNA according to any one of the above (1) to (30) .

(35) A transformant which is obtained by introducing the DNA according to any one of the above (1) to (30) or the recombinant vector of the above (34) into a host cell.

(36) The transformant according to the above (35) wherein the host cell is an avermectin-producing bacterial strain.

(37) The transformant according to the above (35) or (36) wherein the host cell is *Streptomyces avermitilis* K2038 (FERM BP-2775).

(38) A process for producing avermectin aglycon synthase or an avermectin aglycon synthase domain polypeptide comprising:

culturing the transformant according to any one of the above (35) to (37) in a medium to form and accumulate the enzyme or the domain polypeptide in the culture, and recovering the enzyme or the domain polypeptide from the culture.

(39) A process for producing avermectin aglycon or an altered avermectin aglycon comprising:

culturing the transformant according to any one of the above (35) to (37) in a medium to form and accumulate the avermectin aglycon or the altered avermectin aglycon in the culture, and

recovering the avermectin aglycon or the altered avermectin aglycon from the culture.

(40) A process for producing avermectin or altered avermectin comprising:

culturing the transformant according to any one of the above (35) to (37) in a medium to form and accumulate avermectin aglycon or altered avermectin aglycon in the culture,

glycosylating the avermectin aglycon or altered avermectin glycon, and

recovering avermectin or altered avermectin.

(41) The method according to the above (40) wherein altered avermectin is an avermectin which has been altered from avermectin B1a to avermectin B2a.

(42) An altered avermectin obtainable by the process according to the above (40).

(43) An oligonucleotide having a sequence corresponding to 5 to 60 continuous nucleotides in the nucleotide sequence of the DNA according to the above (1) or (2); or an oligonucleotide having a sequence complementary to the oligonucleotide.

Brief Description of the Drawings

Figure 1 is a restriction enzyme map showing *Bam*HI, *Bgl*II, *Cla*I, *Eco*RI, *Kpn*I, *Mlu*I, *Pst*I, *Stu*I and *Xho*I sites of avermectin aglycon synthase genes, aveAI and aveAII, of *Streptomyces avermitilis*. Each arrow indicates the predicted transcriptional direction of each gene.

Figure 2 shows (1) the chromosomal positions of avermectin aglycon synthase genes and the domain sequences of synthase units, (2) the estimated steps of synthesizing avermectin aglycon, and (3) the structure of 6,8a-seco-6,8a-deoxy-5-oxoAvermectin aglycon synthesized with polyketide synthases, which are the gene products of avermectin aglycon synthase genes aveAI and aveAII, and the positions of lower fatty acids which are incorporated into the skeleton of the compound. In this figure, SU indicates synthase unit, ACP indicates acyl carrier protein, AT indicates an acyltransferase, DH indicates dehydratase, DH* indicates a dehydratase-like domain which is estimated to be inactive, KR indicates β-ketoacyl-ACP reductase, KR* indicates a β–ketoacyl-ACP reductase-like domain which is estimated to be active but is not reflected in the polyketide synthetic reaction, KS indicates β-ketoacyl-ACP synthase, and TE indicates thioesterase.

Detailed Description of the Invention

The present invention will be described in detail below.

The present invention relates to DNA sequences of genes encoding avermectin aglycon synthase, and a process for producing avermectin aglycon, a basic constitutional unit of avermectin. According to the present invention, it becomes possible to produce a novel avermectin-associated compound or a specific component of avermectins by modification of the DNA to make a change to the type and number of carboxylic acids to be taken in, modification reaction after condensation, or any combination thereof.

1. Preparation of the DNA of the present invention

A DNA encoding avermectin aglycon synthase (an avermectin aglycon synthase gene) can be isolated from bacteria belonging to the genus *Streptomyces*, e.g. *Streptomyces avermitilis*.

Examples of a method for isolating an avermectin aglycon synthase gene include the method described in Japanese Published Unexamined Patent Application No. 15391/91, colony hybridization described in *Molecular Cloning*, Second Edition, etc.

Specific examples include a method which comprises: ligating the partially digested chromosomal DNA of *Streptomyces avermitilis* with appropriate restriction enzyme such as *Sau*3AI, to a cosmid vector capable of replicating in *E.coli* cleaved at a unique restriction enzyme site, e.g. the vector digested with *Bam*HI; transforming *E.coli* with the obtained recombinant DNA; and selecting a transformant having the avermectin aglycon synthase gene from the obtained transformant by colony hybridization.

Examples of DNAs obtained by the above method include DNAs having the nucleotide sequences shown in SEQ ID NOS: 1 and 2.

The DNA having the nucleotide sequence of SEQ ID NO: 1 or 2 was found by chance to be a DNA fragment encoding a portion of polyketide synthase, when a gene encoding avermectin B5-O-methyl transferase (aveD) was isolated (*Gene*, 206, 175-180 (1998)), and was obtained by the above method.

Modules, domains and ORFs, which are relevant to the avermectin aglycon synthase genes of the present invention, can be determined by comparing similarity with the sequences of 3 types of polyketide synthase domains of erythromycin (*Nature*, 348, 176-178 (1990), *Science*, 252, 675-679 (1991), or *Eur. J. Biochem.*, 204, 39-49 (1992)).

Figure 1 shows a restriction map of the avermectin aglycon synthase gene regions (aveAI and aveAII) of genomic DNA (~65kbp) of *Streptomyces avermitilis* together with predicted transcription units (arrow).

Polyketide compounds are natural organic compounds having a variety of structures and functions, and the common characteristics of these compounds are that their synthesis is carried out with a multifunctional enzyme called polyketide synthase.

One polyketide synthase has substrate specificity, and catalyzes the extention of a lower fatty acid constitutional unit (which is used in the form of CoA ester of dicarboxylic acid in reactions other than intial reaction), i.e., condensation to make a polyketide carbon chain, and has a catalytic activity and a controlling activity which modify a β-carbonyl group generated from such a reaction.

The condensation reaction, which is a basic reaction in the synthesis of polyketide, needs various catalytic activities including an acyl carrier protein (ACP) activity, a β-ketoacyl-ACP synthase (KS) activity and an acyltransferase (AT) activity.

In many cases, β-carbonyl groups generated by the condensation reaction are modified. However, depending on a module, some β-carbonyl groups may not be modified and may be used for the next condensation reaction.

Catalytic activities associated with the modification of a β-carbonyl group after the condensation reaction include a β-ketoacyl-ACP reductase (KR) activity, a dehydratase (DH) activity and an enoyl reductase (ER) activity. The biosynthesis of a polyketide chain is terminated by cleaving out the polyketide chain from polyketide synthase by action of thioesterase (TE) activity.

All or several of these modification activities act in each condensation process, thereby determining the structure of a final product.

The avermectin aglycon synthase genes (aveAI and aveAII) of *Streptomyces avermitilis* are characterized in that the genes have several open reading frames each of

which comprises one or more repeating units called a module, just as with other known polyketide biosynthetic genes. A module is defined as a gene fragment which encodes activities for a one-time synthesis, i.e., a one-time condensation reaction, and the subsequent various modification reactions of the β - carbonyl group. Each module encodes ACP, KS and AT associated with the condensation reaction in polyketide synthesis, and all or several of KR, DH and ER associated with the modification reaction of the β - carbonyl group. Furthermore, there is also a module which does not have any domain for a modification reaction. A polypeptide encoding such a module is referred to as synthase unit (SU).

Figure 2 shows a biosynthetic pathway of 6,8a-seco-6,8a-deoxy-5-oxo-avermectin aglycon synthesized with avermectin aglycon synthases of *Streptomyces avermitilis*.

It is clear that PKS-1 is associated with initiation reaction, since an initiation module (SUs), differing from other modules, has acyltransferase (AT) activity on the N-terminal side. It is clear that PKS-3 is associated with the final reaction of polyketide, since module 9 (SU9) has a thioesterase (TE) domain.

The determined DNA sequences comprising avermectin aglycon synthase genes derived from *Streptomyces avermitilis* are shown in SEQ ID NOS: 1 and 2. The DNA of the present invention comprises open reading frames (ORFs) encoding respective multifunctional enzymes, and these ORFs are ORFs corresponding to nucleotide Nos. 11 to 11916 and nucleotide Nos. 211971 to 30688 of SEQ ID NO: 1 and nucleotide Nos. 31 to 14643 and nucleotide Nos. 414824 to 31419 of SEQ ID NO: 2. The amino acid sequences of various peptides encoded by these sequences are shown in SEQ ID NOS: 3, 4, 5 and 6.

Each of the above DNAs comprises a module encoding a synthesis unit

having all catalytic activities necessary for a one-time carbon chain extension reaction. These modules are represented as the following nucleotides in SEQ ID NOS: 1 and 2. That is to say, the modules are represented in SEQ ID NO: 1 as,

Initiation Module: 85 to 1353,

Module 1: 1441 to 6180,

Module 2: 6256 to 11658,

Module 3: 12076 to 15147,

Module 4: 15217 to 19938,

Module 5: 20008 to 24690,

Module 6: 24781 to 30309, and,

are represented in SEQ ID NO: 2 as,

Module 7: 100 to 4692,

Module 8: 4771 to 7818,

Module 9: 7906 to 14619,

Module 10: 14935 to 20334,

Module 11: 20413 to 25734,

Module 12: 25810 to 31125.

The amino acid sequences of various synthase units (SU) encoded by these modules are represented as the following amino acids. That is to say, the sequences are represented in SEQ ID NO: 3 as,

Initiation SU: 29 to 451,

SU1: 481 to 2060,

SU2: 2086 to 3886;

in SEQ ID NO: 4 as,

SU3: 36 to 1059,

SU4: 1083 to 2656,

SU5: 2680 to 4240,

SU6: 4271 to 6113;

in SEQ ID NO: 5 as,
SU7: 34 to 1564,
SU8: 1591 to 2606,
SU9: 2636 to 4873; and,

in SEQ ID NO: 6 as,
SU10: 38 to 1837,
SU11: 1864 to 3637,
SU12: 3663 to 5434.

DNAs encoding Avermectin aglycon synthase domains (submodules) are represented as the following nucleotides. That is to say, the DNAs are represented in SEQ ID NO: 1 as,
in Initiation Module,
ATs: 85 to 1032,
ACPs: 1096 to 1353;

in Module 1,
KS1: 1441 to 2742,
AT1: 3148 to 4068,
KR1: 5143 to 5676,
ACP1: 5935 to 6180;

in Module 2,
KS2: 6256 to 7545,
AT2: 7906 to 8829,
DH2: 8947 to 9384,

KR2: 10609 to 11142,

ACP2: 11413 to 11658;

in Module 3,

KS3: 12076 to 13368,

AT3: 13756 to 14694,

ACP3: 14902 to 15147;

in Module 4,

KS4: 15217 to 16506,

AT4: 16917 to 17862,

KR4: 18886 to 19419,

ACP4: 19693 to 19938;

in Module 5,

KS5: 20008 to 21297,

AT5: 21658 to 22584,

KR5: 23602 to 24138,

ACP5: 24445 to 24690;

in Module 6,

KS6: 24781 to 26079,

AT6: 26413 to 27336,

DH6: 27475 to 27894,

KR6: 29227 to 29760,

ACP6: 30064 to 30309; and,

are also represented in SEQ ID NO: 2 as,

in Module 7,

KS7: 100 to 1383,

AT7: 1648 to 2673,

KR7: 3634 to 4188,

ACP7: 4447 to 4692;

in Module 8,

KS8: 4771 to 6060,

AT8: 6322 to 7344,

ACP8: 7573 to 7818;

in Module 9,

KS9: 7906 to 9258,

AT9: 9676 to 10773,

DH9: 10885 to 11289,

KR9: 12547 to 13104,

ACP9: 13378 to 13659,

TE9: 13879 to 14619;

in Module 10,

KS10: 14935 to 16224,

AT10: 16543 to 17565,

DH10: 17689 to 18066,

KR10: 19285 to 19842,

ACP10: 20089 to 20334;

in Module 11,

KS11: 20413 to 21705,

AT11: 21991 to 23019,

DH11: 23149 to 23529,

KR11: 24685 to 25242,

ACP11: 25489 to 25734;

in Module 12,

KS12: 25810 to 27102,

AT12: 27367 to 28392,

DH12: 28516 to 28878,

KR12: 30076 to 30633,

ACP12: 30880 to 31125.

The deduced amino acid sequences of various domains encoded by these submodules are represented as:

in SEQ ID NO: 3,

ATs: 29 to 344,

ACPs: 366 to 451,

KS1: 481 to 914,

AT1: 1050 to 1356,

KR1: 1715 to 1892,

ACP1: 1979 to 2060,

KS2: 2086 to 2515,

DH2: 2983 to 3128,

KR2: 3537 to 3714,

ACP2: 3805 to 3886;

in SEQ ID NO: 4,

KS3: 36 to 466,

AT3: 596 to 908,

ACP3: 978 to 1059,

KS4: 1083 to 1512,

Sub G5

AT4: 1653 to 1964,

KR4: 2306 to 2483,

ACP4: 2575 to 2656,

KS5: 2680 to 3109,

AT5: 32030 to 3538,

KR5: 3878 to 4056,

ACP5: 4159 to 4240,

KS6: 4271 to 4703,

AT6: 4815 to 5122,

DH6: 5753 to 5930,

ACP6: 6032 to 6113;

in SEQ ID NO: 5,

KS7: 34 to 461,

AT7: 550 to 891,

KR7: 1212 to 1396,

ACP7: 1483 to 1564,

KS8: 1591 to 2020,

AT8: 2108 to 2448,

ACP8: 2525 to 2606,

KS9: 2636 to 3086,

AT9: 3226 to 3591,

DH9: 3629 to 3763,

KR9: 4183 to 4363,

ACP9: 4460 to 4553,

TE9: 4627 to 4873; and,

in SEQ ID NO: 6,

KS10: 38 to 467,

AT10: 574 to 914,

DH10: 956 to 1081,

KR10: 1488 to 1673,

ACP10: 1756 to 1837,

KS11: 1864 to 2294,

AT11: 2390 to 2732,

DH11: 2776 to 2902,

KR11: 3288 to 3473,

ACP11: 3556 to 3637,

KS12: 3663 to 4093,

AT12: 4182 to 4523,

DH12: 4565 to 4685,

KR12: 5085 to 5270,

ACP12: 5353 to 5434.

From a comparison of sequence information regarding the known polyketide synthase genes, it was found that a similarity of sequences exists between domains having identical functions. By using such similarity, it becomes possible to predict the domain of a novel polyketide synthase gene.

Sub-a6 In other words, based on the above module, domain and ORF information obtained from DNAs having nucleotide sequences of SEQ ID NO: 1 and 2 derived from *Streptomyces avermitilis*, modules, domains and ORFs, which are relevant to the Avermectin aglycon synthase genes of the present invention derived from other bacteria capable of producing avermectin, can be determined.

Using a DNA having the nucleotide sequence of SEQ ID NO: 1 or 2, an avermectin aglycon synthase gene can be obtained by the following method.

A DNA having the nucleotide sequence of SEQ ID NO: 1 or 2 is digested with appropriate restriction enzymes, the DNA fragment was separated and recovered by the method described in *Molecular Cloning*, Second Edition, and an oligonucleotide consisting of the DNA fragment is used as a probe or primer.

As a probe, the DNA fragment labeled with digoxigenin etc. is preferably used. The DIG labeling & detection kit, which can be purchased from Roche Diagnostic Corp., can be used for labeling with digoxigenin.

A library is prepared from bacteria producing avermectin by genome cloning or cDNA cloning described in *Molecular Cloning*, Second Edition etc.

A clone (or a colony) which is to be cross-hybridized with the probe obtained above is selected from the library, then a plasmid is extracted from the clone by the method described in *Molecular Cloning*, Second Edition, and finally an avermectin aglycon synthase gene can be obtained from the plasmid. In addition, DNAs (i.e. submodules) and modules which encode an avermectin aglycon synthase gene domain can be obtained by the same method.

Otherwise, an avermectin aglycon synthase gene, a submodule and a module can also be obtained by direct PCR amplification, using the above library and primers prepared as above.

In a case where the only partial DNA fragment encoding an avermectin aglycon synthase exists in the plasmid extracted as above, according to standard techniques, a restriction map of the plasmid is prepared by digesting the extracted plasmid with appropriate restriction enzymes such as *Bam*HI.

A restriction enzyme map of a DNA comprising the DNA encoding the entire

avermectin polyketide synthase can be prepared by finding restriction enzyme fragments which commonly exist in several clones and binding up the cloned fragments at overlapping portions, so that the DNA encoding avermectin polyketide synthase can be obtained.

The nucleotide sequence of a DNA encoding avermectin polyketide synthase can be determined by commonly applied nucleotide sequencing analysis, e.g. the dideoxy method (*Proc. Natl. Acad. Sci.* USA, 74, 5463 (1977)), or by the analysis with a DNA sequencing analyzer such as 373A DNA sequencer (Perkin Elmer Corp.).

Specifically, a DNA sequence can be determined by directly using double-stranded plasmid DNA as a template for a cycle sequence reaction, wherein oligonucleotide primers specific for various sequences are used. Alternatively, a DNA sequence can also be determined by: cleaving a DNA fragment into its small fragments; introducing the resulting fragments into bacteria phage M13 at random; preparing an overlapping library, which comprises DNA fragments deleted successively from the termini thereof, using a library or plasmid vector comprising partially overlapping fragments; and subjecting each recombinant DNA fragment to the DNA sequencing using oligonucleotide primers specific to the vector sequence. The fluorescently labeled reactant obtained by the cycle sequence reaction can be analyzed with a DNA sequencer (e.g. Model 4000L, LiCor).

Moreover, based on the nucleotide sequence of the determined DNA, a desired DNA can also be prepared by the chemical synthesis, using a DNA synthesizer (Model 8905, Perceptive BioSystems) etc.

The obtained nucleotide sequence data can be arranged, edited and analyzed using existing software, for example, Genetyx™ (Software Development).

Using the DNAs and DNA fragments of the present invention obtained by the above method, oligonucleotides such as an antisense oligonucleotide and a sense oligonucleotide, which have a portion of the DNA sequence of the present invention, or oligonucleotides comprising an RNA can be prepared according to standard techniques. Alternatively, based on the DNA sequence information obtained as above, these oligonucleotides can also be synthesized with the above-mentioned DNA synthesizer.

Examples of the thus obtained oligonucleotides include a DNA having a sequence corresponding to 5 to 60 continuous nucleotides in the nucleotide sequence of the DNA obtained by the above-mentioned method, or a DNA having a sequence complementary to this DNA. Furthermore, the oligonucleotides of the present invention also include an RNA having a sequence complementary to these DNAs.

Examples of the thus obtained oligonucleotides include a DNA having a sequence corresponding to 5 to 60 continuous nucleotides in the nucleotide sequence of SEQ ID NO: 1 or 2, or a DNA having a sequence complementary to this DNA. Where the oligonucletides are used as sense and antisense primers, from among the above oligonucleotides, it is preferable to apply oligonucleotides wherein melting temperature (Tm) and the number of bases do not significantly differ between both oligonucleotides.

Examples of the thus obtained oligonucleotides include ones having the nucleotide sequences shown in SEQ ID NOS: 9 to 14.

Moreover, the derivatives of these oligonucleotides (hereinafter, also referred to as oligonucleotide derivatives) can also be used as the oligonucleotides of the present invention.

Examples of the oligonucleotide derivatives include: an oligonucleotide derivative obtained by conversion of a phosphodiester phosphate bond into a

phosphorothioate bond in the above-described oligonucleotide; an oligonucleotide derivative obtained by conversion of a phosphodiester bond into a N3'-P5' phosphoamidate bond in the above-described oligonucleotide; an oligonucleotide derivative obtained by conversion of a ribose and a phosphodiester phosphate bond into a peptide nucleic acid bond in the above-described oligonucleotide; an oligonucleotide derivative obtained by substitution of uracil by C-5 propynyl uracil in the above-described oligonucleotide; an oligonucleotide derivative obtained by substitution of uracil by C-5 thiazole uracil in the above-described oligonucleotide; an oligonucleotide derivative obtained by substitution of cytosine by C-5 propynyl cytosine in the above-described oligonucleotide; an oligonucleotide derivative obtained by substitution of cytosine by phenoxazine–modified cytosine in the above-described oligonucleotide; an oligonucleotide derivative obtained by substitution of ribose by 2'-O-propyl ribose in the above-described oligonucleotide; and an oligonucleotide derivative obtained by substitution of ribose by 2'-methoxyethoxy ribose in the above-described oligonucleotide etc. (*Cell Engineering* (*Saibo Kogaku*) 16, 1463 (1997)).

2. Preparation of the polypeptide of the present invention

The polypeptide of the present invention can be produced by using a method described in Molecular Cloning, Second Edition or Current Protocols in Molecular Biology. For example, it can be produced by expressing the DNA of the present invention obtained as described in the above Section 1 in a host cell, according to the following procedure.

Based on the DNA of the present invention, a DNA fragment of an appropriate length containing a region encoding the polypeptide of the present invention can be prepared, if necessary. Further, DNA useful for improving the production efficiency of the polypeptide can be prepared by replacing a nucleotide in the nucleotide sequence of the region encoding the polypeptide so as to make a codon most suitable for expression

in a host cell.

The DNA fragment is inserted at a site downstream of a promoter in an appropriate expression vector to construct a recombinant vector.

The recombinant vector is introduced into a host cell suitable for the expression vector, whereby a transformant producing the polypeptide of the present invention can be obtained.

As a host cell, any bacterial cells, yeast cells, animal cells, insect cells, plant cells etc, that are capable of expressing the desired gene can be used.

As an expression vector, it is possible to use any vector that can autonomously replicate in the above host cells or can be integrated into chromosomes thereof and that contains a promoter at a position appropriate for the transcription of the DNA of the present invention.

When a prokaryote (e.g., a bacterial cell) is used as a host cell, a preferred expression vector for the polypeptide of the present invention may be a recombinant DNA construct that is autonomously replicative in prokaryotes and that comprises a promoter, a ribosome-binding sequence, the DNA of the present invention and a terminator. The vector may further comprise a gene regulating the promoter.

Examples of expression vectors include pBTrp2, pBTac1, pBTac2 (each of which is manufactured by Boehringer Mannheim), pKK233-2 (manufactured by Pharmacia), pGEX (manufactured by Pharmacia), pSE280 (manufactured by Invitrogen), pGEMEX-1 (manufactured by Promega), pQE-8 (manufactured by QIAGEN), pQE-9 (manufactured by QIAGEN), pQE-70 (manufactured by QIAGEN), pQE-60 (manufactured by QIAGEN), pET-3 (manufactured by Novagen), pET-11a

(manufactured by Novagen), pKYP10 (Japanese Published Unexamined Patent Application No. 110600/83), pKYP200 [Agric. Biol. Chem., 48, 669 (1984)], pLSA1 [Agric. Biol. Chem., 53, 277 (1989)], pGEL1 [Proc. Natl. Acad. Sci., USA, 82, 4306 (1985)], pBluescript II SK+ (manufactured by Stratagene), pBluescript II SK(-) (manufactured by Stratagene), pTrS30 [prepared from *E. coli* JM109/pTrS30 (FERM BP-5407)], pTrS32 [prepared from *E. coli* JM109/pTrS32 (FERM BP-5408)], pUC19 [Gene, 33, 103 (1985)], pSTV28 (manufactured by Takara Shuzo Co., Ltd.), pUC118 (manufactured by Takara Shuzo Co., Ltd.), pPA1 (Japanese Published Unexamined Patent Application No. 233798/88), pKC30 (Rosenberg et al., 1983, in "Methods in Enzymology," Vol. 101, pp. 123-138, Academic Press, San Diego), pKK223-3 (manufactured by Pharmacia), pDR540 (manufactured by Pharmacia), pRIT2T (manufactured by Pharmacia), and ptrc99a [Gene, 69, 301 (1988)].

As a promoter, any promoter capable of expressing in host cells, such as *E. coli*, can be used, including promoters derived from *E. coli* or a phage such as *trp* promoter (P*trp*), *lac* promoter (P*lac*), P_L promoter, P_R promoter and P_{SE} promoter, SPO1 promoter, SPO2 promoter, and penP promoter. An artificially designed, modified promoter may also be used, including a promoter obtained by binding two P*trp* promoters in tandem (P*trp* × 2), *tac* promoter, lac T7 promoter, and let I promoter.

It is preferable to use a plasmid having an appropriate distance (e.g., 6-18 bases) between Shine-Dalgarno sequence (i.e., ribosome-binding sequence) and an initiation codon.

A terminator is not necessarily required for expression of the recombinant DNA construct of the present invention, but it is desirably located immediately downstream of a structural gene.

A prokaryote includes a microorganism belonging to *Escherichia*, *Serratia*,

Bacillus, *Brevibacterium*, *Corynebacterium*, *Microbacterium*, *Pseudomonas*, and the like. Specific examples include *Escherichia coli* XL1-Blue, *Escherichia coli* XL2-Blue, *Escherichia coli* DH1, *Escherichia coli* MC1000, *Escherichia coli* KY3276, *Escherichia coli* W1485, *Escherichia coli* JM109, *Escherichia coli* HB101, *Escherichia coli* No.49, *Escherichia coli* W3110, *Escherichia coli* NY49, *Serratia ficaria*, *Serratia fonticola*, *Serratia liquefaciens*, *Serratia marcescens*, *Bacillus subtilis*, *Bacillus amyloliquefaciens*, *Brevibacterium ammoniagenes*, *Brevibacterium immariophilum* ATCC14068, *Brevibacterium saccharolyticum* ATCC14066, *Corynebacterium glutamicum* ATCC13032, *Corynebacterium glutamicum* ATCC14067, *Corynebacterium glutamicum* ATCC13869, *Corynebacterium acetoacidophilum* ATCC13870, *Microbacterium ammoniaphilum* ATCC15354, and *Pseudomonas* sp. D-0110.

Introduction of the recombinant DNA can be carried out by any method for introducing DNA into these host cells: for example, the calcium ion method [Proc. Natl. Acad. Sci. USA, 69, 2110 (1972)], the protoplast method (Japanese Published Unexamined Patent Application No. 248394/88) and electroporation [Nucleic Acids Research., 16, 6127 (1988)].

When a yeast cell is used as a host cell, an expression vector which can be used includes YEp13 (ATCC37115), YEp24 (ATCC37051), YCp50 (ATCC37419), pHS19, pHS15, pG-1, pXT1 (manufactured by Stratagene), pSG5 (manufactured by Stragtagene), pSVK3 (manufactured by Pharmacia), pBPV, pMSG (manufactured by Pharmacia), and pSVL SV40 (manufactured by Pharmacia).

As a promoter, any promoter capable of expressing in yeast cells may be used, including PHO5 promoter, PGK promoter, GAP promoter, ADH promoter, GPD promoter, AOX1 promoter, gal 1 promoter, gal 10 promoter, heat shock polypeptide promoter, MF α 1 promoter, and CUP 1 promoter.

Examples of the host cell include yeast strains belonging to the genus *Saccharomyces*, *Schizosaccharomyces*, *Kluyveromyces*, *Trichosporon*, *Schwanniomyces*, *Pichia* and the like. Specific examples include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces lactis*, *Trichosporon pullulans*, *Schwanniomyces alluvius*, or *Pichia pastoris*.

Introduction of the recombinant DNA can be carried out by any method for introducing DNA into yeast cells: for example, electroporation [Methods in Enzymol., 194, 182 (1990)], the spheroplast method [Proc. Natl. Acad. Sci. USA, 81, 4889 (1984)], the lithium acetate method [J. Bacteriol., 153, 163 (1983)] and the method described in Proc. Natl. Acad. Sci. USA, 75, 1929 (1978).

When an animal cell is used as a host cell, an expression vector which can be used includes pcDNAI, pcDM8 (commercially available from Funakoshi), pAGE107 (Japanese Published Unexamined Patent Application No. 22979/91), pAS3-3 (Japanese Published Unexamined Patent Application No. 227075/90), pCDM8 [Nature, 329, 840 (1987)], pcDNAI/Amp (manufactured by Invitrogen), pREP4 (manufactured by Invitrogen), pAGE103 [J. Biochem, 101, 1307 (1987)], pAGE210, pAMo, and pAMoA.

As a promoter, any promoter capable of expressing in animal cells can be used, including a promoter for immediate early (IE) gene of Cytomegalovirus (CMV), SV40 early promoter or metallothionein promoter, retroviral promoter, heat shock promoter, and SRα promoter. An enhancer for IE gene of Human CMV may also be used together with such a promoter.

Host cells include mouse myeloma cells, rat myeloma cells, mouse hybridoma cells, human Namalwa or Namalwa KJM-1 cells, human fetal kidney cells, human leukemia cells, African green monkey kidney cells, chinese hamster CHO cells, or HBT5637 (Japanese Published Unexamined Patent Application No. 299/88).

Sub-C7

Specific examples include SP2/O, NSO and the like for mouse myeloma cells, YB2/O and the like for rat myeloma cells, HEK293 (ATCC: CRL-1573), 293 and the like for human fetal kidney cells, BALL-1 and the like for human leukemia cells, and COS-1, COS-7 and the like for African green monkey kidney cells.

Introduction of the recombinant DNA can be carried out by any method for introducing DNA into animal cells: for example, electroporation [Cytotechnology, 3, 133 (1990)], the calcium phosphate method (Japanese Published Unexamined Patent Application No. 227075/90), the lipofection method [Proc. Natl. Acad. Sci. USA, 84, 7413 (1987)] and the method described in Virology, 52, 456 (1973).

When an insect cell is used as a host cell, a polypeptide can be expressed by a method described in Baculovirus Expression Vectors, A Laboratory Manual, W. H. Freeman and Company, New York (1992); Current Protocols in Molecular Biology; Molecular Biology, A Laboratory Manual; or Bio/Technology, 6, 47 (1988).

More specifically, a recombinant gene transfer vector and a baculovirus may be co-introduced into insect cells to obtain a recombinant virus in the supernatant from the cultured insect cells. Thereafter, insect cells may further be infected with the resulting recombinant virus to express the polypeptide.

Examples of the gene transfer vector used in the above procedure includes pVL1392, pVL1393, pBlueBacIII (commercially available from Invitrogen, respectively) and the like.

Examples of the baculovirus include Autographa californica nuclear polyhedrosis virus, which infects *Noctuidae* insects, and the like.

Examples of insect cells include *Spodoptera frugiperda* ovarian cells, *Trichoplusia ni* ovarian cells, cultured cells derived from silk worm ovary, and the like.

Specific examples are Sf9 and Sf21 (Baculovirus Expression Vectors, A Laboratory Manual) for *Spodoptera frugiperda* ovarian cells, High 5 and BTI-TN-5B1-4 (manufactured by Invitrogen) for *Trichoplusia ni* ovarian cells, *Bombyx mori* N4 for cultured cells derived from silk worm ovary, and the like.

Co-introduction of the recombinant gene transfer vector and the baculovirus into insect cells for recombinant virus production can be carried out by the calcium phosphate method (Japanese Published Unexamined Patent Application No. 227075/90) or the lipofection method [Proc. Natl. Acad. Sci. USA, 84, 7413 (1987)].

When a plant cell is used as a host cell, examples of an expression vector include Ti plasmid, tobacco mosaic virus vector, and the like.

As a promoter, any promoter capable of expressing in plant cells can be used, including cauliflower mosaic virus (CaMV) 35S promoter, rice actin 1 promoter, and the like.

Host cells include plant cells such as tobacco, potato, tomato, carrot, soy bean, Brassica, alfalfa, rice, wheat, barley, and the like.

Introduction of the recombinant vector can be carried out by any method for introducing DNA into plant cells: for example, the *Agrobacterium* method (Japanese Published Unexamined Patent Application No. 140885/84, Japanese Published Unexamined Patent Application No. 70080/85, WO94/00977), electroporation (Japanese Published Unexamined Patent Application No. 251887/85), and the particle gun method (Japanese Patent No. 2606856, Japanese Patent No. 2517813).

The gene can be either expressed directly, or expressed as a secreted polypeptide or a fusion polypeptide according to the method as described in Molecular Cloning, Second Edition. Expression in yeast, animal, insect or plant cells can provide a polypeptide with sugar or sugar chain attached thereto.

The polypeptide of the present invention can be produced by culturing the thus obtained transformant in a medium to produce and accumulate the polypeptide of the present invention in the culture, and recovering the polypeptide from the culture.

The transformant of the present invention can be cultured in a medium according to a conventional method used for culturing host cells.

A medium for culturing a transformant derived from a prokaryote host (e.g., *E. coli*) or a eukaryote host (e.g., yeast) may be a natural or synthetic medium insofar as the medium contains a carbon source, a nitrogen source, an inorganic salt etc., which can be assimilated by the organism, and enables the efficient culture of the transformant.

Any carbon source assimilated by the organisms can be used as a carbon source. Illustrative examples include carbohydrates such as glucose, fructose, sucrose, molasses containing them, starch or starch hydrolysate; organic acids such as acetic acid, propionic acid; alcohols such as ethanol, propanol, and the like.

Examples of the nitrogen source which can be used include ammonium salts of various inorganic or organic acids, such as ammonia, ammonium chloride, ammonium sulfate, ammonium acetate, and ammonium phosphate; other nitrogen-containing compounds; and peptone, meat extracts, yeast extracts, corn steep liquor, casein hydrolysate, soy bean meal, soy bean meal hydrolysate, various fermented cells and hydrolysates thereof and the like.

Inorganic salts which can be used include potassium dihydrogen phosphate, dipotassium hydrogen phosphate, magnesium phosphate, magnesium sulfate, sodium chloride, ferrous sulfate, manganese sulfate, copper sulfate, calcium carbonate, and the like.

Culture is carried out under aerobic conditions by shaking culture, submerged spinner culture under aeration, and the like. The culture temperature is preferably from 15 to 40°C, and culturing time is usually from 5 hours to 7 days. During the culture, pH is maintained at 3.0 to 9.0. The pH can be adjusted using an inorganic or organic acid, an alkaline solution, urea, calcium carbonate, ammonia and the like.

Also, if necessary, antibiotics such as ampicillin and tetracycline can be added to a medium during the culturing.

In a case where a microorganism is transformed with an expression vector containing an inducible promoter, the transformant can be cultured in a medium supplemented with an inducer, if necessary. For example, when an expression vector containing *lac* promoter is used for transformation, the transformant may be cultured in a medium supplemented with isopropyl-β-D-thiogalactopyranoside or the like; when an expression vector containing *trp* promoter is used for transformation, the transformant can be cultured in a medium supplemented with indole acrylic acid or the like.

A medium for culturing a transformant obtained using an animal cell as the host includes generally-used media such as RPMI 1640 medium [The Journal of the American Medical Association, 199, 519 (1967)], Eagle's MEM medium [Science, 122, 501 (1952)], DMEM medium [Virology, 8, 396 (1959)], 199 medium [Proceeding of the Society for the Biological Medicine, 73, 1 (1950)] as well as other media to which fetal calf serum or the like has been added to the above media and the like.

Culturing is generally carried out at pH 6 to 8, at a temperature of 25 to 40°C for a period of 1 to 7 days in the presence of 5% CO_2.

Also, if necessary, antibiotics such as kanamycin, penicillin and streptomycin may be added to a medium during the culturing.

A medium for culturing a transformant obtained using an insect cell as the host includes generally-used media such as TNM-FH medium (manufactured by PharMingen), Sf-900 II SFM medium (manufactured by Life Technologies), ExCell 400 and ExCell 405 [both being products of JRH Biosciences], Grace's Insect Medium [Nature, 195, 788 (1962)] or the like.

Culturing is generally carried out at pH 6 to 7, at a temperature of 25 to 30°C for a period of 1 to 5 days.

Also, if necessary, antibiotics such as gentamycin can be added to a medium during the culture.

The transformant obtained using a plant cell as the host can be cultured as a cell or can be allowed to differentiate into plant cell or organ before culture. Examples of the medium for culturing the transformant include a generally used medium such as Murashige and Skoog (MS) medium, White medium, or any one of these media further supplemented with a plant hormone such as auxin or cytokinin.

Culturing is carried out usually at pH 5 to 9, at a temperature of 20 to 40°C for a period of 3 to 60 days.

Also, if necessary, antibiotics such as kanamycin and hygromycin can be added

to a medium during the culturing.

As described above, the polypeptide of the present invention can be produced by culturing a microorganism-, animal cell-, or plant cell-derived transformant carrying a recombinant vector in which a DNA encoding the polypeptide of the present invention has been inserted according to a general manner to produce and accumulate the polypeptide, and then recovering the polypeptide from the culture.

A method for producing the polypeptide of the present invention includes intracellular production in host cells, extracellular secretion by host cells or production on outer membranes of host cells, and the method can be selected depending on the type of host cells to be used and/or the structure of polypeptide to be produced.

If the polypeptide of the present invention is produced in host cells or on outer membranes of host cells, the polypeptide can be efficiently secreted to extracellularly from the host cells by using the method of Paulson et al. [J. Biol. Chem., 264, 17619 (1989)], the method of Lowe et al. [Proc. Natl. Acad. Sci., USA, 86, 8227 (1989), Genes Develop., 4, 1288 (1990)] or methods as described in Japanese Published Unexamined Patent Application No. 336963/93 and PCT WO94/23021.

More specifically, the polypeptide of the present invention can be efficiently secreted from host cells by expressing it in a form with signal peptide using genetic recombination techniques, the signal peptide being added upstream of a portion containing the active site of the polypeptide of the present invention.

Furthermore, the amount of the production can be increased using a gene amplification system using a dihydrofolate reductase gene or the like according to the method described in Japanese Published Unexamined Patent Application No. 227075/90.

Further, animal or plant cells introduced with a gene may be re-differentiated to create an animal individual carrying a transgene (transgenic non-human animal) or a plant individual carrying a transgene (transgenic plant), which may be used for producing the polypeptide of the present invention.

When the transformant is an animal or plant individual, the polypeptide may be obtained by feeding or cultivating the individual in a general manner to produce and accumulate the polypeptide, and then recovering the polypeptide from the animal or plant individual.

The methods for producing the polypeptide of the present invention using an animal individual include a method using an animal obtained by introducing a gene in accordance with known manners as described in American Journal of Clinical Nutrition, 63, 639S (1996); American Journal of Clinical Nutrition, 63, 627S (1996); and Bio/Technology, 9, 830 (1991).

In the case of an animal individual, for example, the polypeptide of the present invention may be obtained by feeding a transgenic non-human animal introduced with a DNA insert encoding the polypeptide of the present invention to produce and accumulate therein the polypeptide, and then recovering the polypeptide from the animal. The polypeptide can be produced and accumulated in the animal's milk (Japanese Published Unexamined Patent Application No. 309192/88), egg, and the like. As a promoter used for this purpose, any promoter can be used so long as it can be expressed in the animal, for example, mammary gland cell-specific promoters such as an α-casein promoter, a β-casein promoter, a β-lactoglobulin promoter and a whey acidic protein promoter being preferred.

The methods for producing the polypeptide of the present invention using a

plant individual include a method cultivating a transgenic plant obtained by introducing a gen encoding the polypeptide of the present invention to produce and accumulate therein the polypeptide in a known manner as described in Tissue Culture (Soshiki Baiyo), 20 (1994); Tissue Culture, 21 (1995); and Trends in Biotechnology, 15, 45 (1997), and then the polypeptide can be recovered from the plant.

For isolation and purification of the polypeptide produced by the transformant of the present invention, conventional methods for the isolation and purification of enzymes can be used.

For example, if the polypeptide of the present invention is expressed in a soluble form in cells, after completion of culturing, the cells are recovered by centrifugation, and suspended in an aqueous buffer and then disrupted with ultrasonic disrupter, French Press, Manton-Gaulin homogenizer, Dynomill or the like, to obtain a cell-free extract.

From the supernatant obtained by centrifuging the cell-free extract, a purified product can be obtained by the general method used for isolating and purifying an enzyme, for example, solvent extraction, salting-out using ammonium sulfate or the like, desalting, precipitation using an organic solvent, anion exchange chromatography using a resin, such as diethylaminoethyl (DEAE)-Sepharose, DIAION HPA-75 (manufactured by Mitsubishi Chemical) or the like, cation exchange chromatography using a resin, such as S-Sepharose FF (manufactured by Pharmacia) or the like, hydrophobic chromatography using a resin, such as butyl sepharose, phenyl sepharose or the like, gel filtration using a molecular sieve, affinity chromatography, chromatofocusing, or electrophoresis, such as isoelectronic focusing or the like, alone or in combination thereof.

When the protein is expressed as an inclusion body in the host cells, the cells are collected in the same manner, disrupted and centrifuged to recover the inclusion body of the protein as the precipitate fraction. Next, the inclusion body of the protein is solubilized with a protein-denaturing agent.

The solubilized protein solution is diluted with or dialyzed against a solution containing no protein-denaturing agent or such a dilute solution containing the protein-denaturing agent at a lower concentration that denaturation of the protein is not caused. Thus, the normal tertiary structure of the protein is reconstituted. After the procedure, a purified product of the protein can be obtained by a purification and isolation method similar to the above.

When the protein of the present invention or its glycosylated-derivative is secreted out of cells, the protein or its derivative can be collected from the culture supernatant.

Namely, the culture supernatant is obtained by treating the culture in a similar manner to the above-mentioned centrifugation or the like. Then, a purified product can be obtained from the supernatant using a purification and isolation method similar to the above.

Dub-a8 Examples of the thus obtained protein include a protein comprising the amino acid sequence represented by SEQ ID NO:1.

Dub-a9 Furthermore, a fusion protein of the protein of the present invention and other protein may be produced, and purified by affinity chromatography using a substance having affinity to the fusion protein. For example, the protein of the present

Sub-C9

invention may be produced as a fusion protein with protein A according to the method of Lowe *et al.* (*Proc. Natl. Acad. Sci. USA*, *86*: 8227 (1989); *Genes Develop.*, *4*: 1288 (1990)), or the method described in Japanese Published Unexamined Patent Application No. 336963/93 or 823021/94, and purified by affinity chromatography using immunoglobulin G.

Moreover, the protein of the present invention may be produced as a fusion protein with Flag peptide, and the fusion protein can be purified by affinity chromatography using an anti-Flag antibody (*Proc. Natl. Acad. Sci., USA*, *86*: 8227 (1989), *Genes Develop.*, *4*: 1288 (1990)). Further purification can be carried out by affinity chromatography using the antibody against the protein *per se*.

Also, based on the information of the thus obtained protein, the protein of the present invention can be produced by the chemical synthesis method, such as Fmoc (fluorenylmethyloxycarbonyl) method, tBoc (t-butyloxycarbonyl) method, or the like. It can also be chemically synthesized using a peptide synthesizer manufactured by Advanced ChemTech, Perkin-Elmer, Pharmacia, Protein Technology Instrument, Synthecell-Vega, PerSeptive, Shimadzu Corporation, or the like.

3. Production of Avermectin aglycon or Avermectin

Avermectin aglycon can be produced by culturing the transformant prepared in the above Section 2, which carries the avermectin aglycon synthase gene or a module or submodule thereof, in a medium to produce and accumulate avermectin aglycon in the culture, and then recovering avermectin aglycon from the culture.

When a host used for preparation of a transformant can produce avermectin, avermectin aglycon or avermectin can be efficiently produced in any one of the transformants obtained by introducing the avermectin aglycon synthase gene or a

module or submodule thereof into the host. The transformant thus obtained can produce avermectin aglycon or avermectin with higher efficiency than that of the host.

When a host used for preparation of a transformant cannot produce avermectin, the avermectin aglycon synthase gene may be introduced into the host to obtain a transformant capable of producing avermectin aglycon.

In the production of avermectin or avermectin aglycon, the above transformant can be cultured according to a culture procedure as described in the above Section 2.

A known avermectin is a macrocyclic lactone having a 16-membered ring with two sugar residues attached thereto via glycosidic linkage. Avermectin aglycon can be converted into avermectin in a manner well known in the art, for example, by glycosylating avermectin aglycon as described in J. Bacteriol., 175, 2552-2563 (1993).

4. Production of modified Avermectin aglycon or Avermectin

Avermectin aglycon is formed through extension of lower-fatty acid units (used in the form of CoA ester of dicarboxylic acid in reactions other than the initial reaction) by avermectin aglycon synthase, i.e., condensation to give a polyketide carbon chain, and modification of β-carbonyl groups generated during the condensation.

As described above, the avermectin aglycon synthase gene is composed of modules, each module comprising DNA encoding avermectin aglycon synthase domains (submodule).

A submodule encodes ACP, KS and AT involved in condensation during polyketide synthesis, as well as KR, DH or ER involved in modification of β-carbonyl groups.

Accordingly, the carbon chain length of the aglycon part and the type of functional group on β-carbon in the condensation process can be altered by modifying a submodule, based on nucleotide sequence information of the Avermectin aglycon synthase gene determined in the above Section 1.

Further, selective inactivation of a submodule can result in production of a predictable novel avermectin or a particular component thereof alone.

By way of example, a strain producing avermectin B1a and B2a, *Streptomyces avermitilis* K2038, can be converted into a strain producing only avermectin B2a by replacing or converting a submodule DH2 region of the avermectin aglycon synthase gene by or into its inactivated form without dehydratase activity.

The submodule DH2 region can be replaced by or converted into its inactivated form, for example, by homologous recombination on submodule DH2 of the above strain to give the nucleotide sequence shown in SEQ ID NO: 7 in a general manner as described in Molecular Cloning, Second Edition.

The strain thus obtained, which becomes capable of producing a modified (or altered) avermectin aglycon, can be used to produce and obtain the modified avermectin aglycon or modified avermectin according to the general process for producing avermectin.

Best Modes for Carrying Out the Invention

The following examples are provided for illustrative purposes only, and are not intended to limit the scope of the invention clearly defined above.

Example 1: <u>Determination of the nucleotide sequence of the Avermectin aglycon synthase gene of *Streptomyces avermitilis*</u>

A nucleotide sequence of the DNA encoding avermectin aglycon synthase derived from *Streptomyces avermitilis* K2033 (US Patent No. 5206155, FERM BP-2773) was determined as follows.

Continuous or overlapping DNA fragments within the avermectin aglycon synthase gene were subcloned from plasmids containing a fragment of the avermectin aglycon synthase genes (aveAI and aveAII) co-isolated with a gene encoding avermectin B5-O-transmethylase (aveD; Gene, 206, 175-180 (1998)). Nucleotide sequences of the inserted DNA fragments in these subclones were then determined.

More specifically, the entire nucleotide sequences of aveAI and aveAII were determined by subcloning BamHI-treated fragments of 3.4 kbp, 2.0 kbp, 0.5 kbp, 6.8 kbp, 7.0 kbp, 7.8 kbp, 3.7 kbp, 4.8 kbp, 1.3 kbp, 2.4 kbp, 0.7 kbp, 1.0 kbp, 5.4 kbp, 2.5 kbp, 1.9 kbp, 0.1 kbp, 7.0 kbp, 3.1 kbp, 4.7 kbp and 1.3 kbp found in the BamHI-restriction map of aveAI and aveAII shown in Figure 1; treating the inserted DNA fragments in these subclones with exonuclease III and S1 nuclease to prepare a series of deletion fragments; and then carrying out a cycle-sequencing reaction using fluorescently-labeled primers to determine a nucleotide sequence of each deletion fragment. The genes aveAI and aveAII had the nucleotide sequences shown in SEQ ID NO: 1 and SEQ ID NO: 2, respectively.

Example 2: Production of Avermectin B2a alone by nucleotide modification of dehydratase domain in module 2

Streptomyces avermitilis K2033 (FERM BP-2775) produces avermectin B1a and B2a.

The information of avermectin aglycon synthase genes (SEQ ID NOs: 1 and 2) derived from the avermectin-producing strain, which were obtained and sequenced in Example 1, indicated that avermectin was biosynthesized through the biosynthetic

pathway shown in Figure 2.

Avermectin B1a differs from avermectin B2a only in that avermectin B1a has a double bond between 22- and 23-positions in its aglycon part, while avermectin B2a has a single bond between 22- and 23-positions and a hydroxyl group at 23-position in its aglycon part.

The biosynthetic pathway for avermectin found above indicated that avermectin aglycon synthase domain DH of SU2 (DH2) is responsible for the formation of a double bond between 22- and 23-positions in the aglycon part of avermectin B1a and that avermectin aglycon synthase domain DH of SU10 (DH10) has no dehydratase activity.

Based on the idea that a strain producing only avermectin B2a could be obtained by converting DH2 into a domain without dehydratase activity like DH10, DH2 without dehydratase activity was prepared as follows.

The Avermectin aglycon synthase genes obtained in Example 1 had six DH domains: DH2, DH6, DH9, DH10, DH11 and DH12. Comparisons of nucleotide sequences of these domains showed that DH10 clearly differed from other DH domains in a consensus sequence common to dehydratase.

More specifically, amino acid sequences of DH2, DH6, DH9, DH11 and DH12, which were directly involved in the avermectin aglycon formation, were shown to include a consensus sequence of HXaaXaaXaaGXaaXaaXaaXaaP (SEQ ID NO: 13) or HXaaXaaXaaGXaaXaaXaaXaaS (SEQ ID NO: 14), wherein H, G, P, S and Xaa represent histidine, glycine, proline, serine and any amino acid, respectively; whereas the corresponding sequence of DH10 was YXaaXaaXaaGXaaXaaXaaXaaS (SEQ ID NO: 15), wherein Y represents tyrosine, and Xaa, G and S are as defined above.

Thus, replacement of an N-terminal H (His) by Y (Tyr) was thought to provide dehydratase without its activity.

A DNA was constructed, which had a nucleotide sequence encoding the sequence for replacement of the N-terminal amino acid sequence -His-Ala- in the consensus sequence of DH2 [-His-Ala-Val-Gly-Gly-Thr-Val-Leu-Leu-Ser- (SEQ ID NO: 16), amino acids 3037-3046 in SEQ ID NO: 3] by the corresponding sequence of DH10: -Tyr-Glu- (amino acids 1008-1017 in SEQ ID NO: 6). That is, the DNA sequence of DH2 domain: 5'-CAT GCC-3' (nucleotides 9109-9114 in SEQ ID NO: 1) was replaced by the sequence: 5'-TAC GAG-3' as follows.

A DNA fragment of the aveAI region containing DH2 domain was digested with restriction enzyme *Sma*I, and the 2327 bp *Sma*I fragment corresponding to nucleotides 7869-10196 in SEQ ID NO: 1 was cloned into the *Sma*I site of vector plasmid pUC19.

Taq DNA polymerase buffer, dATP, dGTP, dCTP, dTTp and Taq DNA polymerase were added to the resulting recombinant plasmid, which was then divided into two aliquots.

To one of these two aliquots, a primer having the nucleotide sequence shown in SEQ ID NO: 9 (corresponding to nucleotides 9098-9127 in SEQ ID NO: 1) and an antisense primer having the nucleotide sequence shown in SEQ ID NO: 10 (corresponding to an antisense of nucleotides 9193-9222 in SEQ ID NO: 1) were added.

To the other aliquot, an antisense primer having the nucleotide sequence shown in SEQ ID NO: 11 (corresponding to an antisense of nucleotides 9098-9127 in SEQ ID NO: 1) and a primer having the nucleotide sequence shown in SEQ ID NO: 12

(corresponding to nucleotides 8948-8977 in SEQ ID NO: 1) were added.

After the addition, each aliquot was treated at 96 °C for 5 minutes, and the reaction was repeated for 5 to 10 cycles under the following conditions: at 98 °C for 15 seconds and 68 °C for 30 seconds per cycle.

After the reaction, exonuclease I and alkaline phosphatase were added to each aliquot, incubated at 37 °C for 15 minutes, and then treated at 80 °C for 10 minutes to inactivate both the enzymes.

After the inactivation of both the enzymes, Taq DNA polymerase buffer, dATP, dGTP, dCTP, dTTp, a primer having the nucleotide sequence of SEQ ID NO: 12 (corresponding to nucleotides 8948-8977 in SEQ ID NO: 1), an antisense primer having the nucleotide sequence of SEQ ID NO: 10 (corresponding to an antisense of nucleotides 9193-9222 in SEQ ID NO: 1) and Taq DNA polymerase were added to each reaction solution. Each reaction solution was then treated at 96 °C for 5 minutes, and the reaction was repeated for 25 cycles under the following conditions: at 98 °C for 15 seconds and 68 °C for 30 seconds per cycle.

After the reaction, exonuclease I and alkaline phosphatase were added to each reaction solution, incubated at 37 °C for 15 minutes, and then treated at 80 °C for 10 minutes to inactivate the enzymes.

After the inactivation of the enzymes, restriction enzymes *Xcm*I and *Bsa*AI were added to each reaction solution to obtain a *Xcm*I-*Bsa*AI treated DNA fragment.

Restriction enzymes *Xcm*I and *Bsa*AI were added to the recombinant plasmid prepared above, which carried the inserted 2327 bp *Sma*I fragment, to obtain a *Xcm*I-*Bsa*AI treated vector fragment. The *Xcm*I-*Bsa*AI treated vector fragment, T4

DNA ligase buffer, ATP and T4 DNA ligase were added to the *Xcm*I-*Bsa*AI treated DNA fragment, and then incubated overnight at 22 °C to ligate these fragments together, thereby obtaining a plasmid carrying the inserted *Xcm*I-*Bsa*AI treated DNA fragment.

After the transformation of *E. coli* cells with the plasmid, the recombinant plasmids were extracted from individual colonies, and each DNA fragment inserted into the vector was then confirmed for its nucleotide sequence, thereby selecting a clone carrying a fragment introduced with the intended nucleotide replacement.

The inserted DNA fragment was taken from the selected clone, and then carried out recombination with DH2 region on the chromosome of *Streptomyces avermitilis* K2038 by homologous recombination according to a method as described in Japanese Published Examined Patent Application No. 344605/92.

The resulting recombinant *Streptomyces avermitilis* K2210 was cultured under the conditions for general avermectin production, and then the resulting cells were extracted with methanol.

The resulting extract was analyzed using two procedures presented below.

(1) Procedure using thin-layer chromatography on silica gel

Chromatography condition: silica gel, Merck Silica Gel plate F254 (Merck Corp.)

Development solution: n-hexane/iso-propyl alcohol = 85/15

Detection: UV

(2) Procedure using high performance liquid chromatography

Chromatography condition: column, ODS-Hypersil-3 (Elmer Corp.)

Mobile phase: acetonitrile/methanol/water = 60/14/26

Flow rate: 0.6 ml/min

Detection: 246 nm

Temperature: room temperature

In both analytical procedures, only the same peak as that of avermectin B2a was observed. Further, the culture extract was purified by chromatographies on Sephadex LH-20 and silica gel to give the purified product. The purified product was analyzed by NMR and mass spectrometry, indicating that the above recombinant strain produced only avermectin B2a.

Namely, avermectin B2a alone could be produced and obtained according to the method as described above.

Industrial Applicability

The present invention can provide the DNAs encoding a multifunctional enzyme involved in the biosynthesis of avermectin compound useful as a pharmaceutical agent, a veterinary agent and a agricultural chemical; polypeptides encoded by the DNAs; vectors comprising the DNAs; a host cell transformed with the DNA or vector; and a process for producing avermectin or modified avermectin.

Sequence Listing Free Text

SEQ ID NO: 9 represents a primer based on the sequence between nucleotides 9098 and 9127 in SEQ ID NO: 1

SEQ ID NO: 10 represents an antisense primer based on the sequence between nucleotides 9193 and 9222 in SEQ ID NO: 1

SEQ ID NO: 11 represents an antisense primer based on the sequence between nucleotides 9098 and 9127 in SEQ ID NO: 1

SEQ ID NO: 12 represents a primer based on the sequence between nucleotides 8948 and 8977 in SEQ ID NO: 1

The scope of the present invention will be defined by the appended claims, and

it will be appreciated that other numerous variations and modifications may be made without departing from the spirit or scope of the invention. The above examples are therefore to be construed in all respects as illustrative and not restrictive. Further, equivalents of the claims will also fall within the scope of the present invention.

All of patents, patent applications and other publications cited in this specification and the contents as disclosed in the specification and/or drawings of Japanese Patent Application No. 046961/99, which is a priority document of the present application, are incorporated herein by reference in their entirety.